基于无人机高光谱遥感的烤烟叶片叶绿素含量估测

赖佳政，李贝贝，程　翔，孙　丰，陈炬廷，王　晶，张　芊*，叶协锋*

（河南农业大学 烟草学院/国家烟草栽培生理生化研究基地/烟草行业烟草栽培重点实验室，河南郑州 450002）

摘　要：［**目的/意义**］烤烟叶片叶绿素含量（Leaf Chlorophyll Content，LCC）是表征烤烟光合作用、营养状况和长势的重要指标。本研究的目的为高效精确地估测不同生长期烤烟LCC。［**方法**］以中烟100烟叶为研究对象，利用无人机搭载Resonon Pika L高光谱成像仪采集烤烟在6个关键生育期冠层反射率数据。基于相关分析筛选了21种LCC的敏感光谱指数，通过比较不同光谱组合及不同回归分析算法的预测精度，最终建立了基于多种光谱指数组合的LCC回归估测模型。采用一元线性回归（Unary Linear Regression，ULR）、多元线性回归（Multivariable Linear Regression，MLR）、偏最小二乘回归（Partial Least Squares Regression，PLSR）、支持向量回归（Support Vector Regression，SVR）和随机森林回归（Random Forest Regression，RFR）5种建模方法进行LCC估测。［**结果和讨论**］在不同生育期大部分光谱参数与LCC的相关性达到极显著（$P<0.01$）；相较于传统植被指数，新组合的光谱指数显著提升了与LCC的相关性；对单变量LCC估测模型ULR，以移栽后75 d新组合的归一化光谱指数与红光比率光谱指数的单变量建模精度最高，两者决定系数（Coefficient of Determination，R^2）和均方根误差（Root Mean Square Error，RMSE）分别为0.822和0.814，0.226和0.230。MLR、PLSR、SVR和RFR建模方法预测结果表明，RFR算法在LCC估测中效果最好，其中使用移栽后75 d数据验证集的R^2和RMSE可达0.919和0.146。［**结论**］本研究通过分析多种光谱指数与烤烟LCC的响应规律，构建可靠的烤烟叶片LCC估测模型，可为烤烟叶LCC估测以及烤烟的生长发育监测提供理论依据和技术支撑。

关键词：烤烟；叶绿素含量估测；无人机；光谱参数；随机森林回归；多元线性回归；偏最小二乘回归；支持向量机回归

中图分类号：S572;S127　　**文献标志码：**A　　**文章编号：**SA202303007

引用格式：赖佳政, 李贝贝, 程翔, 孙丰, 陈炬廷, 王晶, 张芊, 叶协锋. 基于无人机高光谱遥感的烤烟叶片叶绿素含量估测[J]. 智慧农业(中英文), 2023, 5(2): 68-81.

LAI Jiazheng, LI Beibei, CHENG Xiang, SUN Feng, CHENG Juting, WANG Jing, ZHANG Qian, YE Xiefeng. Monitoring of leaf chlorophyll content in flue-cured tobacco based on hyperspectral remote sensing of unmanned aerial vehicle[J]. Smart Agriculture, 2023, 5(2): 68-81.

1　引　言

叶绿素是作物叶片中主要的光合色素，叶片中叶绿素含量（Leaf Chlorophyll Content，LCC）是反映作物光合作用能力和冠层营养状况的重要指标。烟草作为重要经济作物，以叶片为主要收获对象，因此叶绿素含量监测非常重要。传统的烟草农业监测受田间管理模式及生产设备的限制，费时、费力，效率低下。相较于采用卫星遥感和近地面高光谱人工采集方式，无人机遥感具有分辨率高、效率快、范围广的特点，因此在农作物长势监测方面得以迅速发展和大面积推广应用。

研究表明，基于植被指数（Vegetation Index，VI）的遥感监测具有时空连续性，方便对作物长势信息进行长期监测，在叶绿素估测中具有较高的应

收稿日期：2023-03-14

基金项目：烟草行业烟草栽培重点实验室项目（30800665）；河南省科技攻关项目（172102110168）

作者简介：赖佳政，研究方向为烟草信息学。E-mail：laijiazheng23@163.com

*通信作者：1. 张　芊，博士，讲师，研究方向为烟草信息学。E-mail：Zhangqian225@henau.edu.cn；

2. 叶协锋，博士，教授，研究方向为烟草栽培生理和健康土壤培育。E-mail：yexiefeng@163.com

用价值[1]。冯海宽等[2]通过分析VI与小麦叶绿素相对含量（Soil and Plant Analyzer Development，SPAD）的相关性，将筛选的线性组合指数、植物生化指数、归一化指数以及绿光归一化指数作为偏最小二乘回归（Partial Least Squares Regression，PLSR）模型的输入变量，与冬小麦SPAD值联合构建SPAD-PLSR估测模型，最终确定开花期的SPAD估测精度最好建模集和验证集R^2分别为0.78和0.81。Qiao等[3]研究在不同覆盖度情况下大田玉米的LCC和植被指数的响应特征，采用随机森林（Random Forest，RF）和偏最小二乘两种方法，构建了冠层LCC估测模型。田明璐等[4]选取了28种与作物叶绿素高相关的植被指数，对比发现由多个光谱参数构建的PLSR模型对棉花LCC估测效果最佳。研究人员尝试将高光谱数据对应的波段两两随机组合，形成新的组合光谱指数用以反演作物长势信息。尹航等[5]采用任意两波段组合的光谱指数构建了马铃薯LCC估测模型，相较于其他传统植被指数，由586和498 nm组合的归一化光谱指数以及586和462 nm组合的比率光谱指数构建的反演模型精度和稳定性更为出色。Jia等[6]的研究表明，新组合的光谱指数有助于敏感波段的确定，增强叶片氮含量的相关性，在烤烟叶片氮含量回归模型的性能表现也是可靠的。近年来许多学者结合机器学习算法预测作物农学参数，模型精度和鲁棒性取得较为理想的表现。陈晓凯等[7]通过不同光谱变换下的最优光谱指数结合BP神经网络和RF构建冬小麦叶面积指数的预测模型，RF预测结果优于BP神经网络，并且很好地避免估测模型出现过度拟合和拟合不足的情况。王玉娜等[8]采用随机森林回归（Random Forest Regression，RFR）手段估计了冬小麦的氮营养指数，得到了性能理想的模型（R^2=0.79，RMSE=0.13）。Yang等[9]研究结果显示采用RFR可以显著提高马铃薯块茎形成期和块茎膨大期的地上部生物量估计精度，同时显著降低光谱数据的共线性和冗余度。

目前已发表无人机遥感相关研究多是针对玉米、冬小麦等大田作物，有关烤烟相关研究较少。通过高光谱数据进行烟草冠层LCC的定量分析具有重要意义和应用价值。为实现烤烟LCC的快速估测，本研究利用无人机获取6个不同移栽生育期烤烟叶高光谱影像，通过筛选光谱参数，利用一元线性回归（Unary Linear Regression，ULR）、多元线性回归（Multivariable Linear Regression，MLR）和PLSR三种传统回归方法，支持向量回归（Support Vector Regression，SVR）和RFR两种机器学习建模方法分别构建烤烟的LCC估测模型，并获得最优估测模型，为利用无人机遥感开展田间烤烟长势信息快速、无损监测提供技术支持。

2 材料与方法

2.1 试验材料

试验田位于河南省平顶山市郏县植烟区（112°14′～113°45′E，33°08′～34°20′N），属于温带大陆性季风气候，气候温和，雨量充沛，年日照率53%，平均气温15.7℃，无霜期220天，常年降水量678.5 mm。

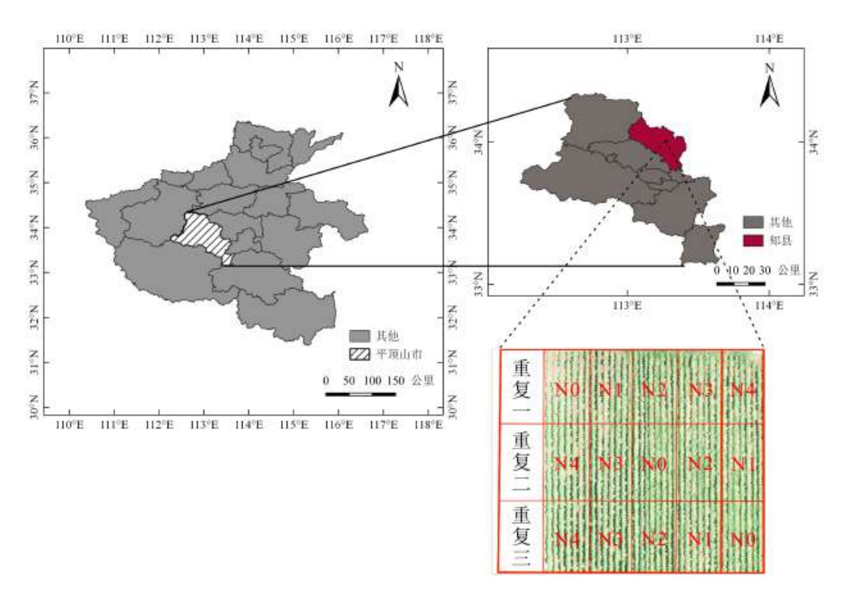
烤烟品种为中烟100。于2022年5月8日移栽，所有肥料均作基肥一次性施入，不再追肥。设置5个水平氮肥梯度，N0～N4处理分别为0.0、24.0、48.0、72.0、96.0 kg/hm²，每个处理设置三次重复，共15个小区，每个小区面积约为144 m²（15.2 m×9.5 m），如图1所示。其他栽培和管理方法按照当地优质烟草栽培方式进行。

2.2 数据获取

2.2.1 无人机高光谱影像获取

获取烤烟6个生育期32、48、61、75、89和109 d的影像。选择中午晴朗、无风、无云的天气条件，采用大疆M600 Pro六旋翼无人机，最大载荷5 kg，最大起飞重量15.5 kg，水平状态下最大的飞行速度能达到18 m/s，续航时间约30 min。搭载Resonon Pika L高光谱成像仪，包含150个波段的光谱数据，主要参数如表1所示。地面试验样品采集与无人机飞机作业同步进行。

试验于上述烤烟生育期11：00—14：00进行测定。无人机飞行作业时，需经过布置在试验地的靶标布便于后期辐射定标处理数据，设置航高100 m，飞行速度3 m/s，最小定时拍照间隔1.0 s，航向重



注：该图由自然资源部标准地图服务网站下载的审图号为GS(2019)1822号标准地图制作，底图无修改

图1 河南省平顶山市烤烟试验地研究区分布图

Fig. 1 Distribution map of the tobacco experimental study area in Pingdingshan city, Henan province

表1 Pika L高光谱成像仪主要参数

Table 1 Main parameters of Pika L hyperspectral imager

参数名称	数值
光谱范围/nm	400～1000
光谱分辨率/nm	2.1
采样间隔/nm	1.07
光谱通道数/个	561
空间通道数/个	900
每秒最大帧数/帧	249
尺寸/cm×cm×cm	10.0×12.5×5.3

叠率80%，旁向重叠率70%，镜头垂直朝下，焦距17 mm。

2.2.2 烤烟LCC的实验室测定

分别获取移栽后32、48、61、75、89和109 d共6个生育期烤烟冠层叶片进行LCC测定。每个小区进行随机破坏性采样3株，每次数据采集共获取45个地面样本。

将烤烟冠层叶片转移到实验室测定LCC。用2 mm打孔器收集新鲜样品，混匀后称量0.1 g碎片，再浸入15 ml的95%乙醇溶液中。在浸提液制备完成后，将其保存在低温、避光和密闭的容器中，在黑暗中提取叶绿素24 h。经过暗处理后，叶子呈白绿色，使用Spectrumlab 722PC比色分光光度计测量叶片色素密度。提取的LCC（Ct）由665和649 nm处的吸光度值计算[10]。LCC计算方法如公式（1）～（3）所示。

$$Ca = 13.95 \times A_{665} - 6.88 \times A_{649} \quad (1)$$

$$Cb = 24.96 \times A_{649} - 7.32 \times A_{665} \quad (2)$$

$$Ct = (Ca+Cb) \times V/1000 \times W \quad (3)$$

其中，Ca和Cb分别表示叶绿素a浓度和叶绿素b浓度，mg/L；V表示浸提液的体积，mL；W表示叶片的质量，g；A_{665}和A_{649}分别为665和649 nm处的吸光度值。

2.3 光谱参数的选定

为了降低土壤背景、大气水分等环境因素对冠层反射率的干扰，大多数研究人员通过利用VI增强作物光谱信息的特征并且实现高光谱数据信息降维[11-13]。本研究将通过两方面选取构建烤烟LCC估测模型的光谱指数：一是基于已发表的对作物LCC敏感的18种植被指数；二是基于波段400～1000 nm范围内任意两个波段随机组合的差值光谱指数（Difference Spectral Index，DSI）、比值光谱指数（Ratio Spectral Index，RSI）和归一化光谱指数（Normalized Spectral Index，NDSI），同时计算3种光谱指数与LCC的相关性并绘制相关性等势图，最终筛选相关系数最优的波段组合为组合光谱指数对LCC进行研究。所选取的光谱参数如表2所示。

表 2 光谱参数和计算公式

Table 2 Spectral indices and formula of calculation

光谱参数	计算公式		参考文献
植物色素比率指数(Plant Pigment Ratio,PPR)	$PPR=(R_{550}-R_{450})/(R_{550}+R_{450})$	(4)	[14]
红边归一化指数(Red edge Normalized Difference Vegetation Index,RNDVI)	$RNDVI=(R_{750}-R_{705})/(R_{750}+R_{705})$	(5)	[15]
红光比率光谱指数(Red Light Ratio Spectral Index,NIR)	$NIR=R_{780}/R_{740}$	(6)	[16]
归一化叶绿素指数(Normalized Difference Chlorophyll Index NDCI)	$NDCI=(R_{762}-R_{527})/(R_{762}+R_{527})$	(7)	[17]
简单比值指数(Simple Ratio Index,SR)	$SR=R_{750}/R_{705}$	(8)	[15]
红光叶绿素光谱指数(Red-edge model index,$CI_{red\text{-}edge}$)	$CI_{red\text{-}edge}=(R_{800}/R_{720})-1$	(9)	[18]
优化土壤调节植被指数(Optimal Soil Adjusted Vegetation Index,OSAVI)	$OSAVI=1.16(R_{800}-R_{670})/(R_{800}+R_{670}+0.16)$	(10)	[19]
红边归一化光谱指数(Normalized Difference Red Edge Index,NDRE)	$NDRE=(R_{790}-R_{720})/(R_{790}+R_{720})$	(11)	[20]
修正归一化光谱指数(Modified Normalized Difference Spectral Index,mND_{705})	$mND_{705}=(R_{750}-R_{705})/(R_{750}+R_{705}+2R_{445})$	(12)	[21]
Vogelmann红边指数(Vogelman Red Edge Index,VOG)	$VOG=R_{740}/R_{720}$	(13)	[22]
红边位置指数(Red Edge Position Index,REP)	$REP=700+40((R_{670}+R_{780})/2-R_{700})/(R_{740}-R_{700})$	(14)	[23]
归一化光谱指数550(ND_{550})	$ND_{550}=(R_{750}-R_{550})/(R_{750}+R_{550})$	(15)	[24]
绿光叶绿素光谱指数(Green Model Index,CI_{green})	$CI_{green}=(R_{800}/R_{560})-1$	(16)	[18]
修正简单比值指数(Modified Simple Ratio Index,mSRI)	$mSRI=(R_{750}-R_{445})/(R_{705}-R_{445})$	(17)	[25]
转换叶绿素吸收反射率指数(Transformed Chlorophyll Absorption in Reflectance Index,TCARI)	$TCARI=3((R_{700}-R_{670})-0.2(R_{700}-R_{550})(R_{700}/R_{670}))$	(18)	[26]
TCARI/OSAVI	$TCARI/OSAVI=[3((R_{700}-R_{670})-0.2(R_{700}-R_{550})(R_{700}/R_{670}))]/[1.16(R_{800}-R_{670})/(R_{800}+R_{670}+0.16)]$	(19)	[26]
绿光归一化植被指数(Green Normalized Difference Vegetation Index,GNDVI)	$GNDVI=(R_{790}-R_{550})/(R_{790}+R_{550})$	(20)	[27]
陆地叶绿素指数(Meris Terrestrial Chlorophyll Index,MTCI)	$MTCI=(R_{754}-R_{709})/(R_{709}-R_{681})$	(21)	[28]
差值植被指数(Difference Vegetation Index,DSI)	$DSI=R_i-R_j$	(22)	本文
比值植被指数(Ratio Spectral Index,RSI)	$RSI=R_i/R_j$	(23)	本文
归一化植被指数(Normalized Difference Spectral Index,NDSI)	$NDSI=(R_i-R_j)/(R_i+R_j)$	(24)	本文

注：i 和 j 表示 400～1000 nm 范围内的任意波段位置；R_i 和 R_j 表示在 i 和 j 位置的反射率

2.4 模型的构建与检验

为了在保证模型训练充分的同时，尽可能减少模型过拟合的风险，以建模集和验证集样本比例2∶1、每个生育期45个样本根据比例随机划分烤烟样本数据。首先对选定的光谱参数和LCC进行相关性分析以筛选对LCC敏感的光谱指数，然后将筛选出的光谱参数作为自变量。参考其他无人机高光谱估测作物长势信息所选择的模型及常见机器学习模型的优势，采用一元线性回归（Unary Linear Regression，ULR）、MLR和PLSR这三种传统线性模型，以及SVR和RFR两种机器学习方法对单个生育期和全生育期烤烟LCC进行预估。

采用决定系数（Coefficient of Determination，R^2）和均方根误差（Root Mean Square Error，RMSE）对模型进行精确性和鲁棒性检验。

3 结果与分析

3.1 烤烟LCC变化

表3描述的是烤烟冠层LCC在不同生育期的变化情况（实验室测定结果）。结果表明，烤烟生长前期，冠层LCC会迅速增加，在团棵期或旺长期LCC达到顶峰；烤烟进入成熟期后，冠层叶片氮素发生转移，LCC开始减少，而烤烟冠层LCC在现蕾期会进一步降低。从整体来看，冠层LCC符合先增后减的变化趋势，在移栽后75 d（旺长期）LCC达到最大，之后开始逐渐下降。在各生育期中移栽后89 d变异系数最大，为28.43%，其余不同生育期的变异系数相差不大，范围为20.38%～25.26%，各个生育期的变异系数均低于全生育期。

全生育期样本范围为0.52～2.95 mg/g，其标准差和变异系数的数值表明LCC的离散程度较大，反

映出试验地不同处理肥力之间是有差异的，也保证了估测模型在一定范围内适用。

表 3 不同生育期烤烟冠层叶片LCC的变化

Table 3 Variation of leaf chlorophyll content（LCC）in flue-cured tobacco leaves at different growth stages

生长期	样本数/个	LCC范围/（mg.g^{-1}）	LCC平均值/(mg.g^{-1})	标准差	变异系数/%
移栽后32 d	45	0.73～2.19	1.53	0.35	23.20
移栽后48 d	45	0.84～2.23	1.61	0.33	20.38
移栽后61 d	45	1.03～2.64	1.80	0.40	21.97
移栽后75 d	45	0.87～2.95	2.05	0.52	25.26
移栽后89 d	45	0.79～2.46	1.60	0.45	28.43
移栽后109 d	45	0.52～1.49	0.88	0.20	22.13
全生育期	270	0.52～2.95	1.58	0.52	33.22

注：全生育期数据集包含6个生育期所有样本；变异系数=标准偏差/平均值

3.2 烤烟LCC和光谱参数的相关性分析

3.2.1 植被指数与烤烟LCC的相关性

将6个生育期的烤烟LCC与植被指数进行相关性分析，结果如表4所示。

由表4可知，除移栽后109 d，其余生育期大部分植被指数和LCC呈极显著相关（$p<0.01$），且大部分植被指数与叶绿素是正相关。其中，移栽后32 d，除TCARI和PPR表现出不相关，其余均为极显著相关，最大相关系数是0.771，为光谱指数NDRE；移栽后48 d，TCARI和PPR为无显著相关，TCARI/OSAVI呈显著相关（$p<0.05$），其余为极显著相关，mND_{705}作为相关性最优的植被指数，其相关系数为0.779；移栽后61 d至移栽后89 d，这三个生育期的植被指数均为极显著相关，PPR、TCARI和TCARI/OSAVI表现为极显著负相关，但三个生育期相关性最高的植被指数各不相同，移栽后61 d相关性最高为mND_{705}，移栽后75 d相关性最高为NIR，移栽后89 d相关性最高为TCARI/OSAVI，其相关系数分别为0.870、0.902和0.877；移栽后109 d，除了PPR、TCARI和TCARI/OSAVI表现为极显著负相关，REP呈显著相关，其余植被指数无显著相关。全生育期样本中所有植被指数和LCC极显著相关，VOG的相关性最大，其相关系数为0.771。此外，在各个生育期红边位置的植被指数

表4 不同生育期烤烟LCC与植被指数的相关系数

Table 4 Correlation coefficients between leaf chlorophyll content(LCC) in flue-cured tobacco and vegetation index at different growth stages

植被指数	移栽后32 d	移栽后48 d	移栽后61 d	移栽后75 d	移栽后89 d	移栽后109 d	全生育期
NIR	0.587**	0.525**	0.698**	0.902**	0.822**	0.250	0.415**
GNDVI	0.672**	0.713**	0.798**	0.848**	0.854**	0.184	0.426**
NDRE	0.771**	0.675**	0.821**	0.883**	0.765**	0.249	0.731**
ND_{550}	0.623**	0.734**	0.783**	0.831**	0.852**	0.179	0.426**
mSRI	0.677**	0.744**	0.768**	0.835**	0.808**	0.229	0.708**
PPR	0.068	−0.032	−0.565**	−0.616**	−0.757**	−0.624**	−0.323**
REP	0.587**	0.530**	0.746**	0.876**	0.849**	0.306*	0.664**
SR	0.698**	0.782**	0.858**	0.852**	0.808**	0.183	0.601**
RNDVI	0.698**	0.792**	0.864**	0.840**	0.827**	0.212	0.530**
VOG	0.738**	0.673**	0.775**	0.869**	0.694**	0.221	0.771**
NDCI	0.482**	0.649**	0.697**	0.742**	0.847**	0.069	0.273**
CI_{re}	0.767**	0.674**	0.817**	0.896**	0.754**	0.219	0.729**
CI_{green}	0.655**	0.725**	0.821**	0.864**	0.830**	0.157	0.490**
mND_{705}	0.685**	0.799**	0.870**	0.846**	0.821**	0.170	0.460**
TCARI	−0.249	0.087	−0.675**	−0.709**	−0.873**	−0.473**	-0.620**
TCARI/OSAVI	−0.590**	−0.311*	−0.759**	−0.761**	−0.877**	−0.404**	-0.630**
OSAVI	0.526**	0.751**	0.467**	0.387**	0.413**	−0.262	0.161**
MTCI	0.691**	0.657**	0.806**	0.866**	0.810**	0.259	0.682**

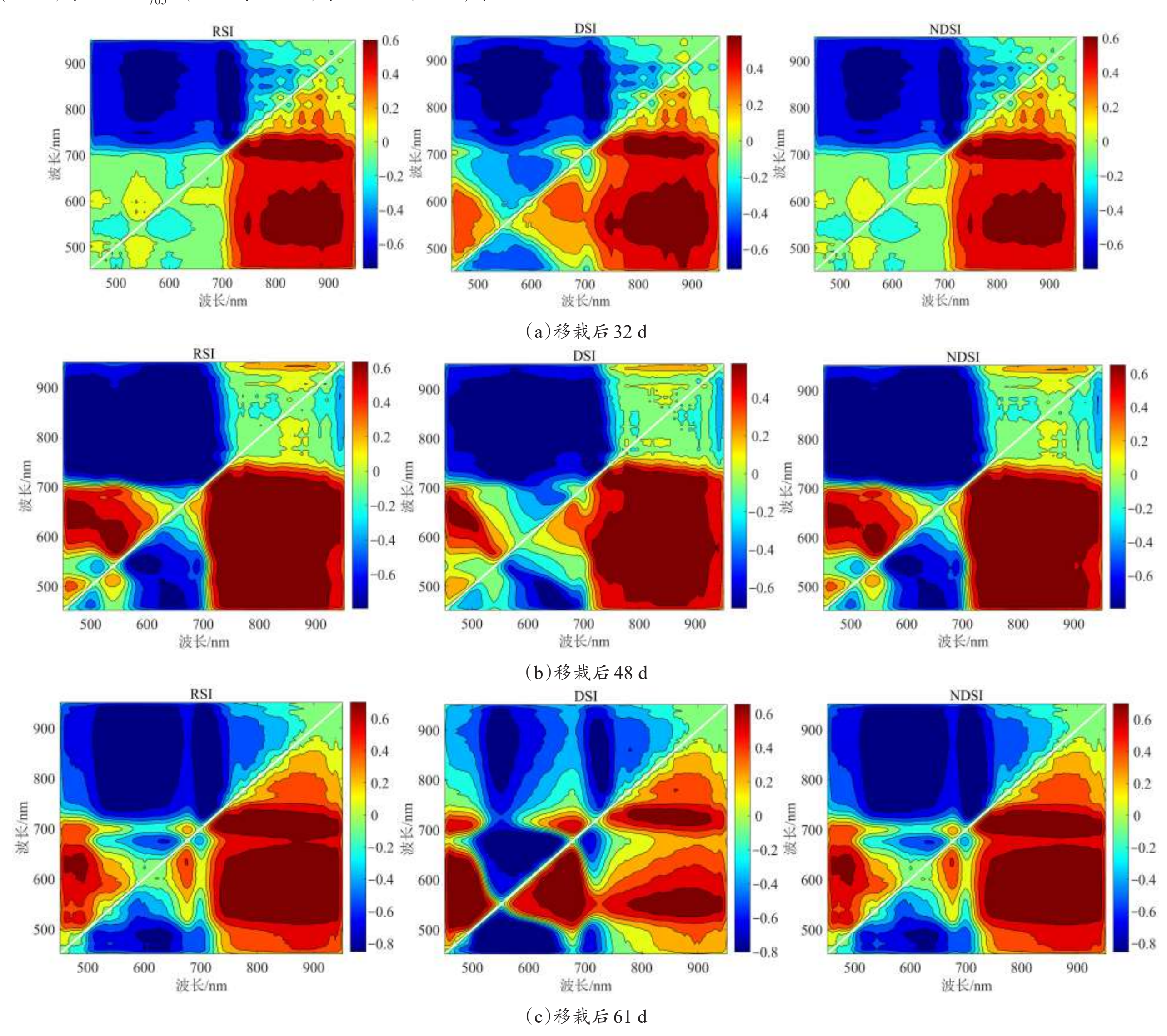
注：* 表示 0. 05 水平显著，**表示 0. 01 水平显著。下同

和LCC含量的相关性均优于其他植被指数，其中NDRE、CIre、RNDVI、VOG以及mND_{705}在不同生育期的相关性，除了移栽后109 d，相关系数均在0.67以上，均呈极显著相关。以上结果表明，VI对LCC的敏感性与作物生长阶段密切相关。随着生育期推进，叶片中的叶绿素会发生转移或降解，这意味着有些VIs不能用于估测LCC。因此，确定不同生长阶段VIs对LCC的敏感性对于基于无人机高光谱图像的作物参数估计非常重要。

对全生育期而言，相关系数绝对值较大的植被指数里绝大部分是红边位置反射率的植被指数。在不同生育期相关性最优的指数分别为NDRE（32 d）、mND_{705}（48 d、61 d）、NIR（75 d）、TCARI/OSAVI（89 d）、TCARI（10 d）以及VOG（全生育期）。分析原因可能是因为NDRE、mND_{705}、NIR和VOG主要使用红边位置和近红外波段信息，这类植被指数对于烤烟LCC的建模估测具有较好的应用潜力，参与建模可以有效提升模型的精确度和鲁棒性[29]。

3.2.2 组合光谱指数与烤烟LCC的相关性

以无人机高光谱波段范围400～1000 nm内任意两波段组合计算得到的光谱指数RSI、DSI和NDSI，并与叶片叶绿素进行相关性分析。计算并绘制各个生育期和全生育期的相关性等势图，如图2和图3所示。



(a)移栽后32 d

(b)移栽后48 d

(c)移栽后61 d

(d)移栽后 75 d

(e)移栽后 89 d

(f)移栽后 109 d

图 2 不同生育期烤烟 LCC 和组合光谱指数的相关性等势图

Fig. 2 Isopotential plots based on correlation between LCC in flue-cured tobacco at different growth stages and new combination of spectral indices

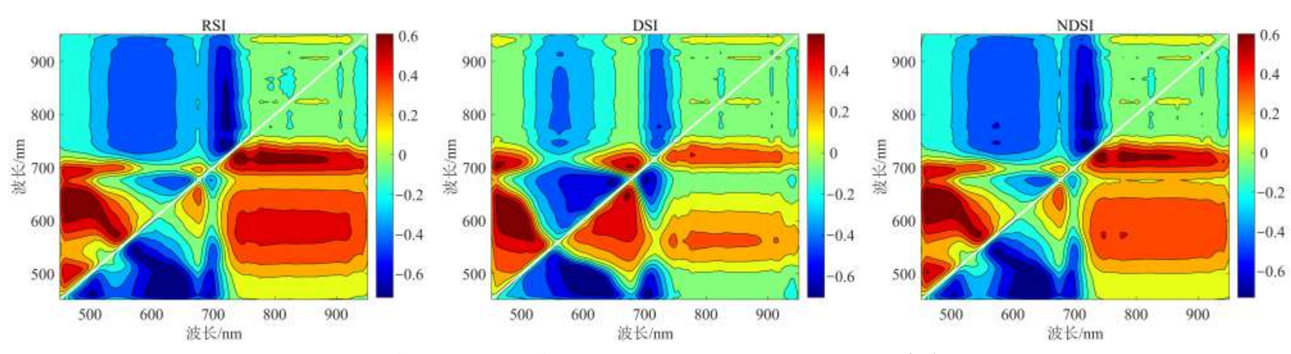


图 3 基于全生育期烤烟 LCC 和组合光谱指数的相关性等势图

Fig. 3 Isopotential plots based on correlation between LCC in flue-cured tobacco leaves during the whole growth stages and new combination of spectral indices

结果表明，除了移栽后109 d，各生育期的敏感波段相对较为集中，与LCC相关性较高的光谱指数组合主要分布在780~940 nm和520~710 nm之间，移栽后109 d相较其他生育期效果不佳。三种随机组合的光谱指数大部分小于相关系数0.3801，达到极显著相关水平的区域较少。全生育期的敏感波段相对较为分散，不同光谱指数之间规律相差不大，但相较于传统植被指数，其相关性有很大提升。以光谱指数与烤烟LCC相关性最高为原则，筛选出不同生育期和全生育期相关系数最优的21种组合光谱指数，如表5所示。结果表明，相较于已有的植被指数，组合光谱指数和LCC的相关性有着更进一步的提升，并且烤烟LCC与筛选出的最优光谱指数均通过0.01水平的显著性检验。对比三种组合光谱指数，归一化光谱指数与烤烟LCC在移栽后32、48、61、75和109 d相关性最高，这些生育期的最优波段组合为797和719 nm、727和697 nm、736和706 nm、775和745 nm以及455和541 nm，其相关系数分别为0.795、0.805、0.878、0.912和0.669；组合差值光谱指数在移栽后89 d时与LCC相关性最高，$DSI_{(R587,\ R579)}$与烤烟LCC的相关系数为0.892；比值光谱指数在全生育期数据集中相关系数最大，$RSI_{(R736,\ R723)}$与烤烟LCC的相关系数为0.807。

表5 烤烟不同移栽生育期组合光谱指数入选波段和LCC相关系数

Table 5 Selected bands of new combination of spectral indices and LCC correlation coefficients at different growth stages

移栽天数/d	光谱指数	相关系数	移栽天数/d	光谱指数	相关系数
32	$DSI_{(R840,R719)}$	0.754**	89	$DSI_{(R587,R579)}$	0.892**
	$RSI_{(R797,R719)}$	0.792**		$RSI_{(R946,R529)}$	0.870**
	$NDSI_{(R797,R719)}$	0.795**		$NDSI_{(R946,R529)}$	0.881**
48	$DSI_{(R766,R732)}$	0.725**	109	$DSI_{(R504,R710)}$	0.521**
	$RSI_{(R719,R697)}$	0.797**		$RSI_{(R455,R541)}$	0.668**
	$NDSI_{(R727,R697)}$	0.805**		$NDSI_{(R455,R541)}$	0.669**
61	$DSI_{(R500,R592)}$	0.822**	全生育期	$DSI_{(R492,R617)}$	0.726**
	$RSI_{(R736,R706)}$	0.878**		$RSI_{(R736,R723)}$	0.807**
	$NDSI_{(R736,R706)}$	0.878**		$NDSI_{(R736,R723)}$	0.806**
75	$DSI_{(R784,R749)}$	0.835**			
	$RSI_{(R775,R745)}$	0.912**			
	$NDSI_{(R775,R745)}$	0.912**			

3.3 烤烟LCC估测模型的构建

3.3.1 基于单个自变量烤烟LCC估测模型

在叶片叶绿素和光谱参数相关性分析的基础上，筛选出6个生育期和全生育期相关系数最高的植被指数NDRE、MND_{705}、NIR、TCARI/OSAVI、PPR和VOG，光谱指数$NDSI_{(R797,\ R719)}$、$NDSI_{(R727,\ R697)}$、$NDSI_{(R736,\ R706)}$、$NDSI_{(R775,\ R745)}$、$DSI_{(R587,\ R579)}$、$RSI_{(R455,\ R541)}$和$RSI_{(R736,\ R723)}$，分别与烤烟叶片叶绿素建立ULR模型，建模集和验证集的估测结果如表6所示。各个生育期的组合光谱指数无论是建模集还是验证集，其估测精度大多优于传统植被，但不同生育期的光谱参数的表现存在差异。植被指数和优化光谱指数从移栽后32 d至移栽后75 d，其建模集和验证集的精度是逐渐增加的，移栽后75 d的NIR作为各个生育期估测效果最佳的植被指数模型，其建模集和验证集的R^2分别是0.814和0.829，RMSE分别是0.230和0.253，而移栽后75 d的$NDSI_{(R775,\ R745)}$是各个生育期估测精度最佳的组合光谱指数，其建模集和验证集的R^2分别是0.822和0.862，RMSE分别是0.226和0.227；移栽后75 d至移栽后109 d建模集和验证集的精度开始下降，而在移栽后109 d时，植被指数和优化组合光谱指数在各个生育期中估测LCC的效果都是最差。相较于其他单个生育期，由不同生育期270个样本组成的全生育期数据集估测表现较为不理想，VOG建模集和验证集的R^2分别是0.602和

0.558，RMSE 分别是 0.348 和 0.349，RSI $_{(R736，R723)}$ 建模集和验证集的 R^2 分别是 0.636 和 0.686，RMSE 分别是 0.333 和 0.304。

表 6 不同生育期单个光谱参数的烤烟LCC估测模型及验证

Table 6 Prediction models and validation of flue-cured tobacco LCC with single spectral parameter at different growth stages

移栽天数/d	光谱参数	回归方程		建模集		验证集	
				R^2	RMSE	R^2	RMSE
32	NDRE	y=10.528x−0.968	（25）	0.606	0.234	0.527	0.198
	NDSI$_{(R797,R719)}$	y=10.912x−1.035	（26）	0.617	0.231	0.662	0.167
48	mND705	y=8.592x−1.811	（27）	0.632	0.193	0.701	0.174
	NDSI$_{(R727,R697)}$	y=9.709x−3.949	（28）	0.646	0.189	0.738	0.163
61	mND705	y=12.592x−3.554	（29）	0.769	0.218	0.707	0.243
	NDSI$_{(R736,R706)}$	y=12.001x−3.651	（30）	0.774	0.216	0.749	0.225
75	NIR	y=17.635x−17.935	（31）	0.814	0.230	0.829	0.253
	NDSI$_{(R775,R745)}$	y=63.199x−0.391	（32）	0.822	0.226	0.862	0.227
89	TCARI/OSAVI	y=−5.264x+3.223	（33）	0.784	0.208	0.772	0.207
	DSI$_{(R587,R579)}$	y=126.691x+4.004	（34）	0.772	0.214	0.847	0.169
109	PPR	y=−5.501x+4.504	（35）	0.413	0.139	0.392	0.169
	RSI$_{(R455,R541)}$	y=7.509x−0.890	（36）	0.473	0.131	0.462	0.149
全生育期	VOG	y=3.124x−3.588	（37）	0.602	0.348	0.588	0.349
	RSI$_{(R736,R723)}$	y=7.026x−7.873	（38）	0.636	0.333	0.686	0.304

3.3.2 基于多自变量烤烟LCC估测模型

基于高光谱遥感的作物生理生化参数估测的研究显示，相较于构建单一光谱参数估测模型，以多个光谱参数作为自变量的建模方法对估测模型精度和鲁棒性都有较为明显的提升[30, 31]。但由于自变量之间存在共线性的问题，在构建烤烟LCC估测模型过程中需要选取适当数量的自变量来防止预测结果过度拟合。因此，本研究选择每个生长生育期以及全生育期数据集中相关系数最高的7个植被指数和3个组合光谱指数作为MLR、PLSR、SVR和RFR四种建模方法的自变量。各个生长阶段通过不同的建模分析方式构建烤烟LCC估测模型，模型的建模集和验证集估测效果如表7所示。

从表7的结果来看，相较于单个光谱参数，多个光谱参数的模型估测精度在各个生长阶段和全生育数据集中有着大幅度的提升。对比各个生育期的建模集估测精度发现，在移栽后32、48、61和109 d，模型LCC-RFR的估测性能最好，R^2分别为0.733、0.744、0.866和0.678，RMSE分别为0.186、0.165、0.176和0.120，而基于LCC-MLR的估测模型，在移栽后75和89 d效果最优，R^2和RMSE分别为0.893、0.850和0.214、0.217。

对比验证集估测精度来看，移栽后32 d，模型LCC-MLR的估测精度最佳，R^2和RMSE分别为0.679和0.215，其余五个生育期模型LCC-RFR的估测精度最佳，R^2分别为0.757、0.832、0.919、0.842和0.580，RMSE分别为0.173、0.185、0.146、0.246和0.131。在全生育数据集中LCC-RFR模型的估测精度较优于LCC-MLR、LCC-PLSR和LCC-SVR模型，对比建模集，R^2增加了19.06%、18.62%和29.51%，RMSE降低了31.93%、29.51%和28.24%；对比验证集，R^2增加了8.21%、12.62%和8.17%，RMSE降低了3.76%、9.33%和4.55%。

综上所述，通过各个生育期的建模集和验证集的估测精度对比分析，LCC-RFR模型的估测效果最为理想，LCC-MLR和LCC-SVR模型估测精度次之，LCC-PLSR模型估测精度最差；不同生育期模型预测精度和稳定性由高到低分别为移栽后75、61、89、48、32和109 d。

3.3.3 烤烟LCC空间分布图

根据不同生育期构建的烤烟LCC估测模型结果可知，LCC-RFR估测模型的精度和稳定性相对最

表7 不同生育期多个光谱参数的烤烟LCC估测模型

Table 7 LCC prediction models of flue-cured tobacco with multiple spectral parameters at different growth stages

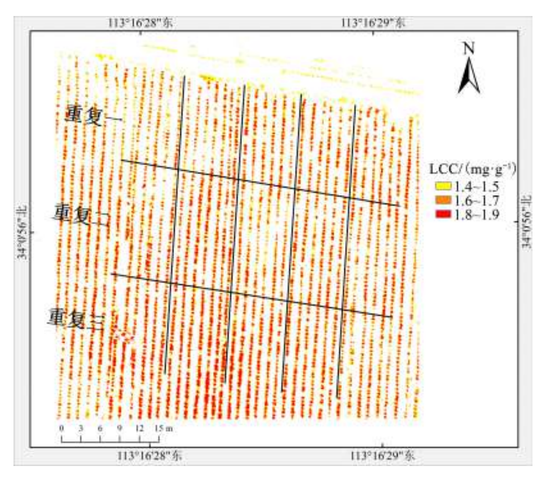
移栽天数/d	模型	建模集		验证集	
		R^2	RMSE	R^2	RMSE
32	LCC-MLR	0.673	0.254	0.679	0.215
	LCC-PLSR	0.590	0.247	0.626	0.162
	LCC-SVR	0.647	0.232	0.569	0.176
	LCC-RFR	0.733	0.186	0.660	0.210
48	LCC-MLR	0.680	0.240	0.742	0.148
	LCC-PLSR	0.607	0.217	0.621	0.192
	LCC-SVR	0.702	0.680	0.691	0.164
	LCC-RFR	0.744	0.165	0.757	0.173
61	LCC-MLR	0.812	0.232	0.820	0.264
	LCC-PLSR	0.746	0.238	0.750	0.227
	LCC-SVR	0.830	0.197	0.736	0.228
	LCC-RFR	0.866	0.176	0.832	0.185
75	LCC-MLR	0.893	0.214	0.804	0.289
	LCC-PLSR	0.766	0.293	0.807	0.204
	LCC-SVR	0.856	0.204	0.824	0.269
	LCC-RFR	0.891	0.205	0.919	0.146
89	LCC-MLR	0.850	0.217	0.781	0.224
	LCC-PLSR	0.765	0.221	0.799	0.224
	LCC-SVR	0.790	0.218	0.810	0.184
	LCC-RFR	0.846	0.163	0.842	0.246
109	LCC-MLR	0.631	0.134	0.511	0.181
	LCC-PLSR	0.484	0.119	0.429	0.192
	LCC-SVR	0.513	0.117	0.606	0.168
	LCC-RFR	0.678	0.120	0.580	0.131
全生育期	LCC-MLR	0.718	0.303	0.741	0.274
	LCC-PLSR	0.720	0.292	0.712	0.291
	LCC-SVR	0.731	0.287	0.742	0.276
	LCC-RFR	0.854	0.206	0.802	0.264

佳。因此，本研究在消除土壤背景和杂草的干扰后，利用LCC-RFR估测模型选择烤烟三个生育期进行无人机高光谱影像的反演填图，得到了移栽后48、61和75 d的烤烟LCC空间分布图，如图4所示。整体来看，从移栽后48到75 d，随着生育期的推进颜色逐渐加深，表明烤烟LCC越来越大。不同重复之间也有所差异，烤烟长势分布由强到弱为“重复三”“重复二”“重复一”，对于移栽后45 d，LCC的范围在14～1.9 mg/g之间，从移栽后61 d开始，重复一与重复二、重复三LCC的差异愈发明显，重复一LCC估测的范围大多在1.8～2.3 mg/g，重复二和重复三估测LCC大多在2.1～2.5 mg/g范围内。移栽后75 d LCC整体分布较密，LCC的范围在2.2～3.1 mg/g之间，这可能与烤烟生长特性有关，在烤烟生育前期至旺长期，烤烟LCC呈逐渐上升趋势。通过估测多个生育期的LCC，在无人机高光谱图像上反演的LCC空间分布与实测结果相符，这表明无人机搭载高光谱成像仪可提供小范围内烤烟LCC的空间分布信息，具有较高的分辨率和精度，可以较好地分辨出烤烟各个生育期烤烟长势区别。

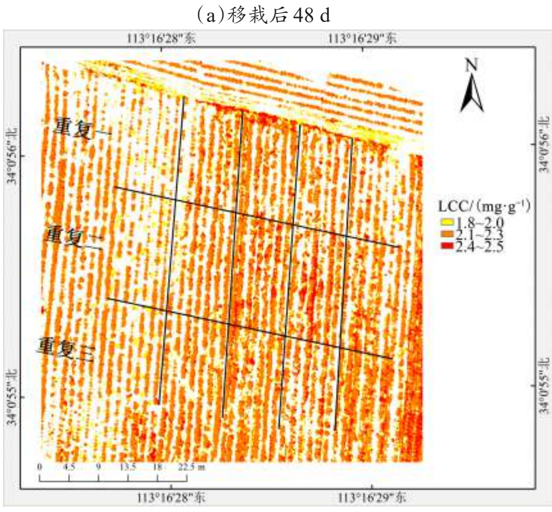
4 讨 论

本研究通过将烤烟LCC与选取的光谱参数进行相关分析，结果表明本研究构建的组合光谱指数与LCC的相关性明显高于传统植被指数，且基于单个光谱构建LCC回归模型中，前者的估测精度和稳定性在各个生育期包括全生育期都优于后者。前人研究表明构建组合光谱指数能够有效减少外界环境因素干扰，精准识别作物生长参数的敏感波段，对农作物氮素、叶绿素和叶面积指数的监测精度以及模型鲁棒性有显著提升[5,7,11]。但也有学者认为，组合光谱指数所包含的光谱信息不足，不能充分利用从可见光到近红外波段的高光谱数据信息[32]。因此，本研究采用传统植被指数和优化光谱指数相结合的方式试图构建更为综合的烤烟LCC估测模型。试验结果表明，基于多个光谱参数组合构建的烤烟LCC估测模型精度和拟合性都优于单个光谱参数构建的模型。多个自变量可以有效避免光谱信息包含不足的情况，充分利用高光谱数据，但是光谱参数过多也容易出现共线性和过拟合的问题，因此，今后研究需要关注对烤烟LCC估测模型自变量的控制。

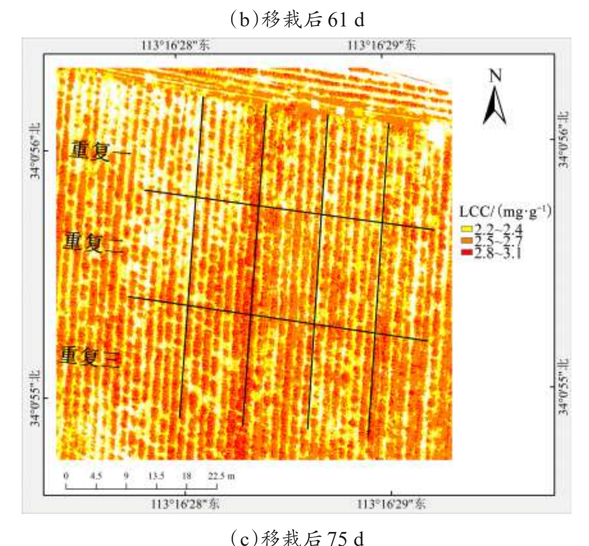
此外，不同生育期的模型精度各不相同，且随着生育期的推进，模型精度呈现先增后减的变化趋势，这可能是由于烤烟在生长发育的不同阶段，伴随着冠层结构的变化、生物量的积累，以及在生殖生长阶段冠层叶片氮含量的分配，导致光谱参数的预测性能在不同阶段发生变化。对比分析MLR、



(a)移栽后48 d



(b)移栽后61 d



(c)移栽后75 d

图4 烤烟不同生育期的LCC空间分布图

Fig. 4 LCC spatial distribution map of flue-cured tobacco plants at different growth stages

PLSR、SVR和RFR四种多自变量LCC估测模型，基于RFR模型估测烤烟LCC的精度和稳定性效果最好，主要原因是RFR建模时不易出现过拟合现象，能够有效降低异常值和噪声的干扰[33]。利用多生育期数据集构建全生育期模型会降低模型预测精度[34]，一方面可能是因为冠层结构和土壤背景情况随着生育期的推进在不断改变，由冠层光谱对LCC的响应规律发生变化造成，另一方面，可能是由不同生育期冠层光谱的测定误差叠加造成。此外，目前本研究的数据集涵盖了不同生育阶段和氮素水平下的样本，但构建的组合光谱指数尚未在不同环境和年份中得到验证。为了提高模型的稳健性和泛化能力，需要进一步测试不同烤烟品种的适用性。

5 结 论

本研究探索豫中烟区烤烟LCC与无人机高光谱影像的响应规律，通过多种光谱指数与烤烟LCC的相关分析，构建烤烟叶片叶绿素估测含量的ULR、MLR、PLSR、SVR和RFR模型，得到以下结论：

（1）所涉大部分植被指数和LCC呈极显著相关。其中，移栽后32—109 d相关系数最大的植被指数分别是NDRE、MND_{705}、mND_{705}、NIR、TCARI/OSAVI和REP，全生育期数据集中所有植被指数与烤烟LCC均呈极显著相关，VOG的相关性最大（r = 0.771）。

（2）根据随机两两波段组合的光谱指数与烤烟LCC含量相关性最大原则，筛选出不同生育期和全生育期相关系数最高的21种光谱指数，传统植被指数和组合光谱指数与LCC的敏感波段大多集中在红边位置，但整体上组合光谱指数的相关性较前者更为出色，其中移栽后75 d的$NDSI_{(R775,\ R745)}$和$RSI_{(R775,\ R745)}$的相关性最高，相关系数r均为0.912。

（3）在叶片叶绿素和光谱参数相关性分析的基础上，筛选出6个生育期和全生育期烤烟LCC相关性最高的光谱参数与烤烟叶片叶绿素分别建立ULR模型，不同移栽生育期的优化光谱指数无论是建模集还是验证集其估测精度均较优于传统植被指数。

（4）相较于单个光谱参数，基于多个光谱参数构建模型的估计精度在各个生长阶段和全生育数据集中均有大幅度提升，其中移栽后75 d由LCC-RFR构建的模型最佳，验证集R^2和RMSE可达

0.919和0.146。综合来看，RFR模型的LCC估测效果最为理想，LCC-MLR和LCC-SVR模型估测精度次之，LCC-PLSR模型估测精度最差；不同生育期模型预测精度和稳定性由高到低分别为移栽后75、61、89、48、32和109 d。

利益冲突声明：本研究不存在研究者以及与公开研究成果有关的利益冲突。

参考文献：

[1] WANG L, CHEN S S, LI D, et al. Estimation of paddy rice nitrogen content and accumulation both at leaf and plant levels from UAV hyperspectral imagery[J]. Remote sensing, 2021, 13(15): ID 2956.

[2] 冯海宽, 陶惠林, 赵钰, 等. 利用无人机高光谱估算冬小麦叶绿素含量[J]. 光谱学与光谱分析, 2022, 42(11): 3575-3580.
FENG H K, TAO H L, ZHAO Y, et al. Estimation of chlorophyll content in winter wheat based on UAV hyperspectral[J]. Spectroscopy and spectral analysis, 2022, 42(11): 3575-3580.

[3] QIAO L, TANG W J, GAO D H, et al. UAV-based chlorophyll content estimation by evaluating vegetation index responses under different crop coverages[J]. Computers and electronics in agriculture, 2022, 196: ID 106775.

[4] 田明璐, 班松涛, 常庆瑞, 等. 基于无人机成像光谱仪数据的棉花叶绿素含量反演[J]. 农业机械学报, 2016, 47(11): 285-293.
TIAN M L, BAN S T, CHANG Q R, et al. Estimation of SPAD value of cotton leaf using hyperspectral images from UAV-based imaging spectroradiometer[J]. Transactions of the Chinese society for agricultural machinery, 2016, 47(11): 285-293.

[5] 尹航, 李斐, 杨海波, 等. 基于无人机高光谱影像的马铃薯叶绿素含量估测[J]. 植物营养与肥料学报, 2021, 27(12): 2184-2195.
YIN H, LI F, YANG H B, et al. Estimation of canopy chlorophyll in potato based on UAV hyperspectral images[J]. Journal of plant nutrition and fertilizers, 2021, 27(12): 2184-2195.

[6] JIA F F, LIU G S, LIU D S, et al. Comparison of different methods for estimating nitrogen concentration in flue-cured tobacco leaves based on hyperspectral reflectance[J]. Field crops research, 2013, 150: 108-114.

[7] 陈晓凯, 李粉玲, 王玉娜, 等. 无人机高光谱遥感估算冬小麦叶面积指数[J]. 农业工程学报, 2020, 36(22): 40-49.
CHEN X K, LI F L, WANG Y N, et al. Estimation of winter wheat leaf area index based on UAV hyperspectral remote sensing[J]. Transactions of the Chinese society of agricultural engineering, 2020, 36(22): 40-49.

[8] 王玉娜, 李粉玲, 王伟东, 等. 基于无人机高光谱的冬小麦氮素营养监测[J]. 农业工程学报, 2020, 36(22): 31-39.
WANG Y N, LI F L, WANG W D, et al. Monitoring of winter wheat nitrogen nutrition based on UAV hyperspectral images[J]. Transactions of the Chinese society of agricultural engineering, 2020, 36(22): 31-39.

[9] YANG H B, LI F, WANG W, et al. Estimating above-ground biomass of potato using random forest and optimized hyperspectral indices[J]. Remote sensing, 2021, 13(12): ID 2339.

[10] 萧浪涛, 王三根. 植物生理学实验技术[M]. 北京: 中国农业出版社, 2005.
XIAO L T, WANG S G. Experimental technology of plant physiology[M]. Beijing: China Agriculture Press, 2005.

[11] LIU H Y, ZHU H C, WANG P. Quantitative modelling for leaf nitrogen content of winter wheat using UAV-based hyperspectral data[J]. International journal of remote sensing, 2017, 38(8-10): 2117-2134.

[12] 付波霖, 邓良超, 张丽, 等. 联合星载高光谱影像和堆栈集成学习回归算法的红树林冠层叶绿素含量遥感反演[J]. 遥感学报, 2022, 26(6): 1182-1205.
FU B L, DENG L C, ZHANG L, et al. Estimation of mangrove canopy chlorophyll content using hyperspectral image and stacking ensemble regression algorithm[J]. National remote sensing bulletin, 2022, 26(6): 1182-1205.

[13] ZHANG Y, XIA C Z, ZHANG X Y, et al. Estimating the maize biomass by crop height and narrowband vegetation indices derived from UAV-based hyperspectral images[J]. Ecological indicators, 2021, 129: ID 107985.

[14] METTERNICHT G. Vegetation indices derived from high-resolution airborne videography for precision crop management[J]. International journal of remote sensing, 2003, 24(14): 2855-2877.

[15] GITELSON A, MERZLYAK M N. Spectral reflectance changes associated with autumn senescence of *Aesculus hippocastanum* L. and *Acer platanoides* L. leaves. spectral features and relation to chlorophyll estimation[J]. Journal of plant physiology, 1994, 143(3): 286-292.

[16] MISTELE B, SCHMIDHALTER U. Tractor-based quadrilateral spectral reflectance measurements to detect biomass and total aerial nitrogen in winter wheat[J]. Agronomy journal, 2010, 102(2): 499-506.

[17] RANJAN R, CHOPRA U K, SAHOO R N, et al. Assessment of plant nitrogen stress in wheat (*Triticum aestivum* L.) through hyperspectral indices[J]. International journal of remote sensing, 2012, 33(20): 6342-6360.

[18] GITELSON A A, GRITZ Y, MERZLYAK M N. Relationships between leaf chlorophyll content and spectral reflectance and algorithms for non-destructive chlorophyll assessment in higher plant leaves[J]. Journal of plant physiology, 2003, 160(3): 271-282.

[19] RONDEAUX G, STEVEN M, BARET F. Optimization of soil-adjusted vegetation indices[J]. Remote sensing of environment, 1996, 55(2): 95-107.

[20] BARNES E, CLARKE T, RICHARDS S E, et al. Coincident detection of crop water stress, nitrogen status and canopy density using ground-based multispectral data[C]//

Proceedings of the 5th International Conference on Precision Agriculture. Madison, USA: American Society of Agronomy, 2000.
[21] SIMS D A, GAMON J A. Relationships between leaf pigment content and spectral reflectance across a wide range of species, leaf structures and developmental stages[J]. Remote sensing of environment, 2002, 81(2/3): 337-354.
[22] VOGELMANN J E, ROCK B N, MOSS D M. Red edge spectral measurements from sugar maple leaves[J]. International journal of remote sensing, 1993, 14(8): 1563-1575.
[23] CLEVERS JAN G P W. Imaging spectrometry in agriculture - plant vitality and yield indicators[M]// Eurocourses: remote sensing. Dordrecht: Springer Netherlands, 2007: 193-219.
[24] GITELSON A A, KAUFMAN Y J, MERZLYAK M N. Use of a green channel in remote sensing of global vegetation from EOS-MODIS[J]. Remote sensing of environment, 1996, 58(3): 289-298.
[25] DATT B. A new reflectance index for remote sensing of chlorophyll content in higher plants: Tests using *Eucalyptus* leaves[J]. Journal of plant physiology, 1999, 154(1): 30-36.
[26] HABOUDANE D, MILLER J R, TREMBLAY N, et al. Integrated narrow-band vegetation indices for prediction of crop chlorophyll content for application to precision agriculture[J]. Remote sensing of environment, 2002, 81(2/3): 416-426.
[27] GITELSON A A, MERZLYAK M N. Remote sensing of chlorophyll concentration in higher plant leaves[J]. Advances in space research, 1998, 22(5): 689-692.
[28] DASH J, CURRAN P J. The *MERIS* terrestrial chlorophyll index[J]. International journal of remote sensing, 2004, 25(23): 5403-5413.
[29] 陈倩, 常庆瑞, 郭松, 等. 基于红边特性和连续小波变换的冬小麦叶绿素含量估算[J]. 麦类作物学报, 2022, 42(7): 883-891.
CHEN Q, CHANG Q R, GUO S, et al. Estimation of chlorophyll content in winter wheat based on red edge characteristics and continuous wavelet transform[J]. Journal of triticeae crops, 2022, 42(7): 883-891.
[30] LI C C, CHEN P, MA C Y, et al. Estimation of potato chlorophyll content using composite hyperspectral index parameters collected by an unmanned aerial vehicle[J]. International journal of remote sensing, 2020, 41(21): 8176-8197.
[31] TAO H L, FENG H K, XU L J, et al. Estimation of crop growth parameters using UAV-based hyperspectral remote sensing data[J]. Sensors, 2020, 20(5): ID 1296.
[32] 班松涛, 田明璐, 常庆瑞, 等. 基于无人机高光谱影像的水稻叶片磷素含量估算[J]. 农业机械学报, 2021, 52(8): 163-171.
BAN S T, TIAN M L, CHANG Q R, et al. Estimation of rice leaf phosphorus content using UAV-based hyperspectral images[J]. Transactions of the Chinese society for agricultural machinery, 2021, 52(8): 163-171.
[33] CHEN X K, LI F L, SHI B T, et al. Estimation of winter wheat canopy chlorophyll content based on canopy spectral transformation and machine learning method[J]. Agronomy, 2023, 13(3): ID 783.
[34] PATEL M K, RYU D, WESTERN A W, et al. Which multispectral indices robustly measure canopy nitrogen across seasons: Lessons from an irrigated pasture crop[J]. Computers and electronics in agriculture, 2021, 182: ID 106000.

Monitoring of Leaf Chlorophyll Content in Flue-Cured Tobacco Based on Hyperspectral Remote Sensing of Unmanned Aerial Vehicle

LAI Jiazheng, LI Beibei, CHENG Xiang, SUN Feng, CHENG Juting, WANG Jing, ZHANG Qian*, YE Xiefeng*

(*College of Tobacco Science, Henan Agricultural University/ National Tobacco Cultivation and Physiology and Biochemistry Research Center/ Key Laboratory for Tobacco Cultivation of Tobacco Industry, Zhengzhou* 450002, *China*)

Abstract:

[Objective] Leaf chlorophyll content (LCC) of flue-cured Tobacco is an important indicator for characterizing the photosynthesis, nutritional status, and growth of the crop. Tobacco is an important economic crop with leaves as the main harvest object, it is crucial to

monitor its LCC. Hyperspectral data can be used for the rapid estimation of LCC in flue-cured tobacco leaves, making it of great significance and application value. The purpose of this study was to efficiently and accurately estimate the LCC of flue-cured tobacco during different growth stages.

[Methods] Zhongyan 100 was chose as the research object, five nitrogen fertilization levels were set. In each plot, three plants were randomly and destructively sampled, resulting in a total of 45 ground samples for each data collection. After transplanting, the reflectance data of the flue-cured tobacco canopy at six growth stages (32, 48, 61, 75, 89, and 109 d) were collected using a UAV equipped with a Resonon Pika L hyperspectral. Spectral indices for the LCC estimation model of flue-cured tobacco were screened in two ways: (1) based on 18 published vegetation indices sensitive to LCC of crop leaves; (2) based on random combinations of any two bands in the wavelength range of 400–1000 nm. The Difference Spectral Index (DSI), Ratio Spectral Index (RSI), and Normalized Spectral Index (NDSI) were calculated and plotted against LCC. The correlations between the three spectral indices and leaf LCC were calculated and plotted using contour maps. Five regression models, unary linear regression (ULR), multivariable linear regression (MLR), partial least squares regression (PLSR), support vector regression (SVR), and random forest regression (RFR), were used to estimate the chlorophyll content. A regression estimate model of LCC based on various combinations of spectral indices was eventually constructed by comparing the prediction accuracies of single spectral index models multiple spectral index models at different growth stages.

[Results and Discussions] The results showed that the LCC range for six growth stages was 0.52–2.95 mg/g. The standard deviation and coefficient of variation values demonstrated a high degree of dispersion in LCC, indicating differences in fertility between different treatments at the test site and ensuring the applicability of the estimation model within a certain range. Except for 109 d after transplanting, most vegetation indices were significantly correlated with LCC ($p<0.01$). Compared with traditional vegetation indices, the newly combined spectral indices significantly improved the correlation with LCC. The sensitive bands at each growth stage were relatively concentrated, and the spectral index combinations got high correlation with LCC were mainly distributed between 780–940 nm and 520–710 nm. The sensitive bands for the whole growth stages were relatively dispersed, and there was little difference in the position of sensitive band between different spectral indices. For the univariate LCC estimation model, the highest modeling accuracy was achieved using the newly combined Normalized Spectral Index and Red Light Ratio Spectral Index at 75 d after transplanting. The coefficients of determination (R^2) and root mean square errors (RMSE) for the modeling and validation sets were 0.822, 0.814, and 0.226, 0.230, respectively. The prediction results of the five resgression models showed that the RFR algorithm based on multivariate data performed best in LCC estimation. The R^2 and RMSE of the modeling set using data at 75 d after transplanting were 0.891 and 0.205, while those of the validation set reached 0.919 and 0.146. In addition, the estimation performance of the univariate model based on the whole growth stages dataset was not ideal, with R^2 of 0.636 and 0.686, and RMSE of 0.333 and 0.304 for the modeling and validation sets, respectively. However, the estimation accuracy of the model based on multiple spectral parameters was significantly improved in the whole growth stages dataset, with R^2 of 0.854 and 0.802, and RMSE of 0.206 and 0.264 for the modeling and validation sets of the LCC-RFR model, respectively. In addition, in the whole growth stages dataset, the estimation accuracy of the LCC-RFR model was better than that of the LCC-MLR, LCC-PLSR, and LCC-SVR models. Compared with the modeling set, R^2 increased by 19.06%, 18.62%, and 29.51%, while RMSE decreased by 31.93%, 29.51%, and 28.24%. Compared with the validation set, R^2 increased by 8.21%, 12.62%, and 8.17%, while RMSE decreased by 3.76%, 9.33%, and 4.55%.

[Conclusions] The sensitivity of vegetation indices (VIs) to LCC is closely connected to the tobacco growth stage, according to the results this study, which examined the reaction patterns of several spectral indices to LCC in flue-cured tobacco. The sensitivity of VIs to LCC at various growth stages is critical for crop parameter assessment using UAV hyperspectral photography. Five estimation models for LCC in flue-cured tobacco leaves were developed, with the LCC-RFR model demonstrating the greatest accuracy and stability. The RFR model is less prone to overfitting and can efficiently decrease outlier and noise interference. This work could provide theoretical and technological references for LCC estimate and flue-cured tobacco growth monitoring.

Key words: flue-cured tobacco; chlorophyll content monitoring; unmanned aerial vehicles; spectral parameters;random forest regression; multivariable linear regression; partial least squares regression; support vector regression